

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 033 401 A2

2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:  
06.09.2000 Bulletin 2000/36

(51) Int Cl.7: **C12N 15/12**, C07K 14/47,  
C07K 16/18, G06F 17/30

(21) Application number: **00200610.4**

(22) Date of filing: **21.02.2000**

(84) Designated Contracting States:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE**  
Designated Extension States:  
**AL LT LV MK RO SI**

(30) Priority: **26.02.1999 US 122487 P**

(71) Applicant: **GENSET**  
**75008 Paris (FR)**

(72) Inventors:  
• **Dumas Milne Edwards, Jean-Baptiste**  
**75006 Paris (FR)**  
• **Duclert, Aymeric**  
**94100 Saint-Maur (FR)**

• **Giordano, Jean-Yves**  
**75018 Paris (FR)**

(74) Representative: **Brasnett, Adrian Hugh et al**  
**Mewburn Ellis**  
**York House**  
**23 Kingsway**  
**London WC2B 6HP (GB)**

Remarks:

THE COMPLETE DOCUMENT INCLUDING  
REFERENCE TABLES AND THE SEQUENCE  
LISTING IS AVAILABLE ON CD-ROM FROM THE  
EUROPEAN PATENT OFFICE, VIENNA  
SUB-OFFICE.

(54) **Expressed sequence tags and encoded human proteins**

(57) The sequences of 5' ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5' ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used

in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.

06-SEP-2000.  
21-FEB-2000; 0000EP-0309610.  
26-FEB-1999; 990S-0122487.  
(GEST ) GENSET.  
Dumas Milne Edwards J, Duclert A. Giordano J;  
WPI: 2000-500381/45.  
N-PSDB: AAC00310.  
New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
diagnostic, forensic, gene therapy and chromosome mapping procedures -  
Claim 13; SEQ ID 4385; 71pp + CD-ROM; English.  
The present sequence is a polypeptide encoded by one of a large number  
of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
were prepared from total human RNAs or polyA+ RNAs derived from 30  
different tissues. EST sequences usually correspond mainly to the 3'  
untranslated region (UTR) of the mRNA because they are often obtained  
from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
those cases where longer cDNA sequences have been obtained, the full 5'  
UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
ends and can therefore be used to obtain full length cDNAs and genomic  
DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
chromosome mapping procedures. They are used to obtain upstream  
regulatory sequences and to design expression and secretion vectors.  
Query Match 20.9%; Score 456; DB 21; Length 83;  
Best Local Similarity 98.8%; Pred. No. 3.5e-35;  
Matches 82; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MARVLGAPVALGSLWLSLAIAIPLPTSAHGNVAEGTKPDQVTERCSGWSFDATT 60  
1 MARVLGAPVALGSLWLSLAIAIPLPTSAHGNVAEGTKPDQVTERCSGWSFDATT 60  
61 LDNCTMLFFKGFYVWVSKHNDP 83  
61 LDNCTMLFFKGFYVWVSKHNDP 83  
RESULT 3  
ABP31577  
ID ABP31577 standard; Protein; 87 AA.  
XX AC ABP31577;  
XX DT 09-JUL-2002 (first entry)  
XX DE Human glycoprotein-like ORF550 protein, SEQ ID NO:1100.  
XX KW Human; ORF: open reading frame; ORF: drug screening; diagnosis;  
XX KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
XX KW immune modulation; haematopoiesis regulation; tissue growth;  
XX KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;  
XX KW thrombolytic; tumour inhibition; bodily characteristics; fertility;  
XX KW behaviour; cancer; proliferative disorder; neurological disorder;  
XX KW cardiovascular disease; immune system disorder; organ transplantation;  
XX KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
XX KW hypothyroidism; cholesterol ester storage disease; infection; vulnerability;  
XX KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; neurotropic;  
XX KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
XX KW cardiatic; hyotensive; antithyroid; antiinflammatory; immunomodulator;  
XX KW dermatological; analgesic; virocid; antibacterial; fungicide.  
XX OS Homo sapiens.

06-SEP-2000.  
21-FEB-2000; 0000EP-0309610.  
26-FEB-1999; 990S-0122487.  
(GEST ) GENSET.  
Dumas Milne Edwards J, Duclert A. Giordano J;  
WPI: 2000-500381/45.  
N-PSDB: AAC00310.  
New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
diagnostic, forensic, gene therapy and chromosome mapping procedures -  
Claim 13; SEQ ID 4385; 71pp + CD-ROM; English.  
The present sequence is a polypeptide encoded by one of a large number  
of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
were prepared from total human RNAs or polyA+ RNAs derived from 30  
different tissues. EST sequences usually correspond mainly to the 3'  
untranslated region (UTR) of the mRNA because they are often obtained  
from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
those cases where longer cDNA sequences have been obtained, the full 5'  
UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
ends and can therefore be used to obtain full length cDNAs and genomic  
DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
chromosome mapping procedures. They are used to obtain upstream  
regulatory sequences and to design expression and secretion vectors.  
Query Match 20.9%; Score 456; DB 21; Length 83;  
Best Local Similarity 98.8%; Pred. No. 3.5e-35;  
Matches 82; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 MARVLGAPVALGSLWLSLAIAIPLPTSAHGNVAEGTKPDQVTERCSGWSFDATT 60  
1 MARVLGAPVALGSLWLSLAIAIPLPTSAHGNVAEGTKPDQVTERCSGWSFDATT 60  
61 LDNCTMLFFKGFYVWVSKHNDP 83  
61 LDNCTMLFFKGFYVWVSKHNDP 83  
RESULT 3  
ABP31577  
ID ABP31577 standard; Protein; 87 AA.  
XX AC ABP31577;  
XX DT 09-JUL-2002 (first entry)  
XX DE Human glycoprotein-like ORF550 protein, SEQ ID NO:1100.  
XX KW Human; ORF: open reading frame; ORF: drug screening; diagnosis;  
XX KW disease monitoring; cytokine; cell proliferation; cell differentiation;  
XX KW immune modulation; haematopoiesis regulation; tissue growth;  
XX KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;  
XX KW thrombolytic; tumour inhibition; bodily characteristics; fertility;  
XX KW behaviour; cancer; proliferative disorder; neurological disorder;  
XX KW cardiovascular disease; immune system disorder; organ transplantation;  
XX KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;  
XX KW hypothyroidism; cholesterol ester storage disease; infection; vulnerability;  
XX KW vasotrophic; antipsoriatic; antidiabetic; cytostatic; neurotropic;  
XX KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;  
XX KW cardiatic; hyotensive; antithyroid; antiinflammatory; immunomodulator;  
XX KW dermatological; analgesic; virocid; antibacterial; fungicide.  
XX OS Homo sapiens.